The LaserROM Project: A Case Study In Document Processing Systems

Mike Rafeld
Application Support Division
Hewlett-Packard

Each year, Hewlett-Packard maintains, updates, and distributes over 8000 unique publications which are used by thousands of customers worldwide. This translates into hundreds of thousands of pages. Documentation distribution is now the province of an entire operation within HP. Because the effective use of HP products depends on the effective use of supporting publications and manuals, HP is constantly trying to improve their documentation distribution. Early in 1987, the compact disc audio (CD-Audio) market was quickly maturing. At the same time, an offshoot of this technology was just beginning, called Compact Disc-Read Only Memory (CD-ROM). The Application Support Division at Hewlett-Packard realized that this new technology for data storage might make a significant impact on the problem of documentation distribution. It could provide large amounts of documentation and support information to HP's customers in a timely, reliable, and consistent manner. This would, in turn, increase the usefulness of HP's documentation, thereby helping the customer make better support decisions and solidifying our relationship with them. Thus, the LaserROM project was created to determine how to put the thousands of publications produced by Hewlett-Packard onto CD-ROM.

CD-ROM was seen as a viable delivery vehicle for a number of reasons. Mastering costs run $3,000 to $5,000 for the first few discs, and then pressing of additional discs costs approximately three to ten dollars per disc.The storage capacity of a CD-ROM is greater than 600 megabytes, which translates to 250,000 typed pages. The data transfer rate is 150 kilobytes per second, which translates into a .4 to 1 second access time [1,2]. CD-ROM publishing technology had progressed to the point where it was economically feasible to update and distribute these discs on a very timely basis, such as once a month.

Three major challenges now faced the LaserROM development team. The first was to design a process to consolidate documents and data which existed in variety of formats and put them onto the compact disc. The second was to design the representation of the data on the CD-ROM. And the third was to design a user interface to retrieve the data. These three issues are the main focus of this paper.

The first section deals with the consolidation and preparation of the data to be put on the compact disc. This process takes a document in its original format from the information providers. The number of different types of formats is large, and this process converges all of these varied formats to the data representation format, which is described in the second section.

This second section discusses the representation of the documentation so that it conforms to the constraints of the LaserROM interface, and how graphic illustrations were dealt with. Finally, the document is converted into the physical format in which it will reside on the compact disc. Then, a magnetic tape is generated, which is then sent to the CD mastering facility. Refer to Figure 1 for a graphical representation of the CD-ROM production process.

The third section deals with the retrieval end of the LaserROM project. Interface issues included defining how the retrieval would be implemented, and how these retrieval mechanisms could be integrated into an easy, intuitive user interface.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and / or specific permission.

© 1988 ACM 0-89791-291-8/88/0012/0021 $1.50

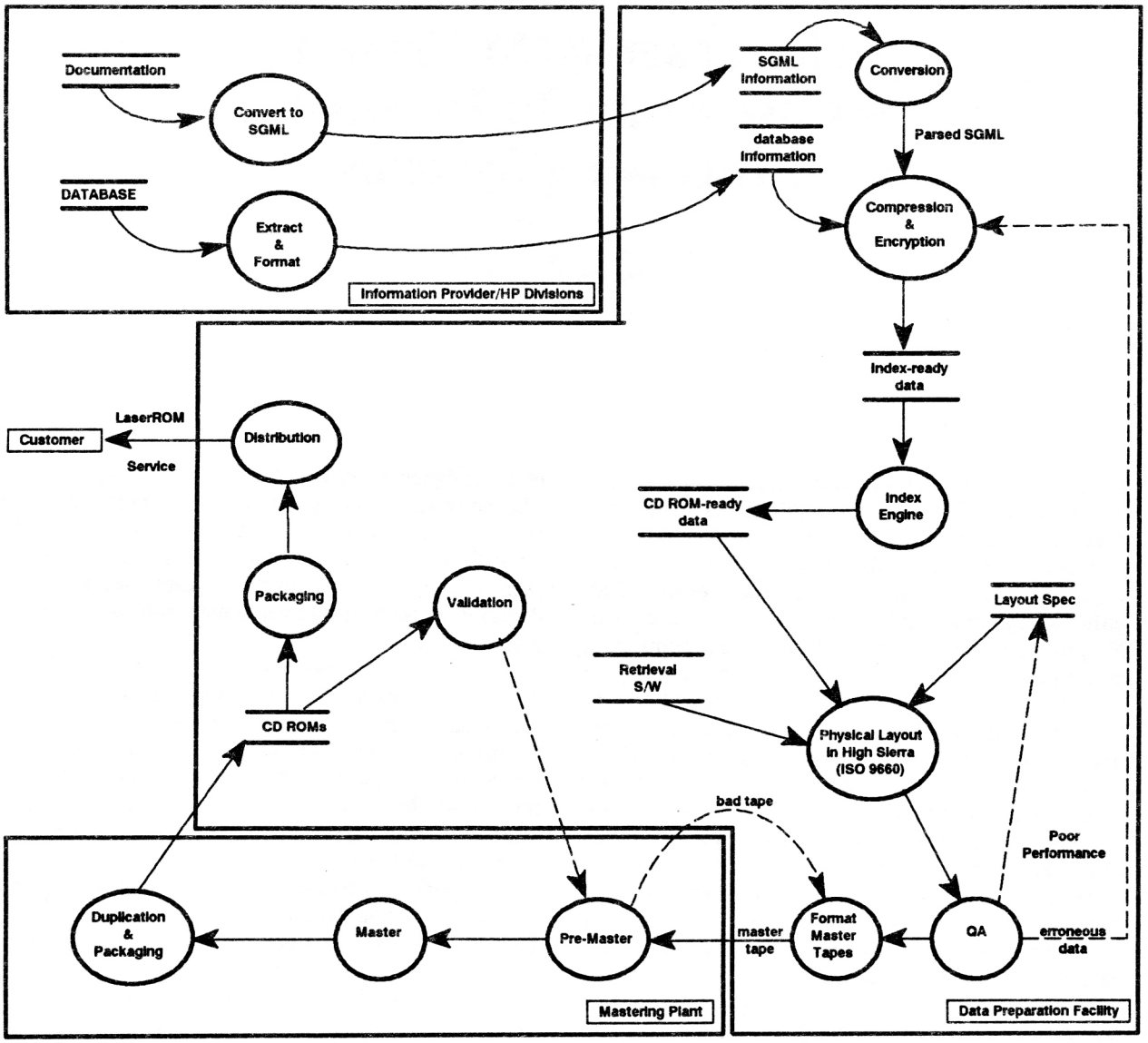

Figure 1. This flowchart illustrates the overall CD-ROM production process, which takes place when data is collected and formatted for the CD-ROM.

In all three of these sections, the use of existing standards is emphasized, and attention is given to how these standards were integrated into the LaserROM product.

Finally, a fourth section explores what's in the future, and how LaserROM might be adapted to meet the demands of changing customer needs or improving technologies. This section also discusses some of the unique problems associated with the production of a CD-ROM, and how these problems were overcome at Hewlett-Packard.

The Data Preparation Process

Selecting A Document Format

User documentation exists in a variety of formats. Some of the technical documentation the LaserROM team was faced with was produced using word processors only available in an HP3000 environment (e.g. TDP, MANU), while others are produced using common publishing systems (e.g. Interleaf, Ventura Publisher), while still others had been sent to the printers as pasted-up documents, with no electronic version of the

manual extant. Important support information also existed not as manuals, but in actual databases, and this data needed to be extracted and formatted, as well. The challenge was to take manuals and data of all these different types and convert them to a consistent format on the compact disc.

With all of this in mind, it was obvious that some common input format should be chosen, and all documents and data which were to be on the LaserROM CD-ROM would need to be submitted in this format. As shown in Figure 1, the information providers providing data to be distributed on LaserROM include the various HP divisions. Each division is responsible for the technical manuals for its products, and, until recently, this meant that there were almost as many different document formats as there were HP divisions.

At this same time, a standard called the Standardized Generalized Markup Language [3] (SGML) was being investigated by Corporate Documentation Support. SGML "is a notation for representing documents and making their inherent structure explicit. Various forms of automatic processing can be performed on documents coded in SGML; they can be formatted, loaded into online databases, or analyzed for various linguistic purposes." [4] SGML is defined in International Standard 8879 (ISO 8879).

The charter of Corporate Documentation Support was to create a markup language based on SGML which would be used in the generation of all future documentation at Hewlett-Packard. The motivation for this was to have the same look for all of HP's technical documentation, and to facilitate the interchange of documents and personnel within the corporation. The language which was created was called HP Tag [5].

Software tools to support HP Tag (and SGML) were provided by Corporate Documentation Support. These tools generalize some constructs of SGML, while omitting other SGML constructs. HP Tag has differences which mostly fall into the category of SGML constructs that have not been implemented, although they may be provided in future releases [4]. The notation defining HP Tag is isomorphic to the SGML document type definition. MARKUP is the tool provided to support use of SGML. It consists of an SGML parser and facilities for defining applications on SGML documents [4,6].

Because SGML was an accepted standard, and HP Tag had corporate and upper-level management support to become the documentation standard within Hewlett-Packard, and so that LaserROM would be able to accommodate future HP documentation, HP Tag was the obvious choice for a document format. A document marked up to use HP Tag might look like Figure 2.

```
< glossary >
< title > Glossary of Animals < \title >

< entry >
< term > Aardvark < \term >
< definition >
The first animal listed in a dictionary.
< \definition >
< \entry >

< entry >
< term > Dog < \term >
< definition >
A domesticated canine, probably descended
from the < xref > wolf < \xref > .
< \definition >
< \entry >

      .
      .
      .
< \glossary >
```

Figure 2. A sample HP Tag document, with all minimization expanded. If the user so chose, this markup could actually be much simpler, since the conventions for markup minimization can determine the remaining structure from a document-type definition which specifies a *glossary* as a *title* followed by *entries* followed by blank lines.

Converting Existing Documentation

The next problem was managing the conversion of literally hundreds of documents which did not exist in HP Tag format. After considering the volume of documentation which needed to be converted, and the time constraints, the LaserROM team decided upon using both in-house expertise as well as external vendors to assist in the conversion effort.

Conversion was achieved in one of three ways: 1) By creating a program which would convert an SGML-like language (e.g. Copy Markup Language (CML), a documentation language internal to HP) to HP Tag, or 2), actually going into the electronic copy with a text editor and marking it up by hand, or 3) for those manuals which had no corresponding electronic form, scanning them in using Optical Character Recognition (OCR) software which converted the information into ASCII format, and then marking it up by hand.

The conversion effort was **the** major undertaking in producing the initial CD-ROM. It took 15 to 20 people, working for three months, using a combination of the methods enumerated above, to convert approximately 30 manuals. Each manual averaged about 200 pages.

To provide some idea of what this effort involved, the programmatic conversion was limited to a CML to HP Tag program, which ran on an IBM PC. This program would take 20 to 30 minutes to convert a 200 page manual, and approximately 70% of the manual was correctly converted. About 30% of each manual needed to be patched up by hand.

Marking up manuals completely by hand was an effort that depended very heavily on the complexity of the manual being converted. An average-sized manual of 200 pages, all text, would take one person four days to convert. And, the same size manual, with half text and half tables, would take one person eight-and-a-half days to convert. A 200-page manual which contained half text, and half illustrations (discussed below), would take seven person days. Experience showed that tables required the most effort to convert from their current format to the SGML-based HP Tag. All of these times are estimated for manuals which already existed in some electronic format—manuals which needed to be scanned in would require the additional time to scan in the information.

Extra time was needed for proofreading of the manuals, once they were marked up. Every manual which was marked up had either text missing, or incorrect tags inserted, or an inordinate number of misspellings. Incorrectly marking up a manual could cause tables or other complex constructs to be generated incorrectly, and scanned text was especially proofread for misspellings.

Integrating Illustrations

HP Tag addressed the textual data, but the issue of illustrations which existed in the manuals was not automatically solved by using HP Tag. How could these graphics be presented to the user, and how could they be integrated seamlessly into the documents as they were being formatted?

The interface to the LaserROM project was being developed at the same time, and its constraints were that it was an MS-Windows-based application which was designed to run on a number of PC platforms. (The interface is discussed in greater detail in the third section.) Rather than integrating the text and graphics on the same screen for the user, since text referring to the illustration might be off-screen already, the LaserROM team decided that figures in documents would be represented as icons on the viewscreen. If the user then wanted to look at the corresponding figure, they would activate the icon by using a pointing device. They could then scroll through the textual data at their leisure, while referring to the illustration as needed.

In terms of document processing, this meant that hooks were needed in the formatted document to account for the figure file references. This was achieved using existing tags in HP Tag. This also meant determining how to store the figures on the CD-ROM, and how they would be displayed on the viewscreen. The display of the graphic illustrations is discussed in the third section.

The storage format chosen was the Tagged Image File Format (TIFF), which is a recognized interchange format for figures, used widely in the desktop publishing community. TIFF was chosen for a number of reasons. Firstly, a number of commercial scanners, chiefly the HP Scanjet, will generate TIFF files. The interface also required that a bit-mapped (raster) format be used so that the graphic output could be portable across a number of platforms.

The TIFF standard was developed by Aldus Corporation and Microsoft Corporation, with the assistance of several vendors. It provides features that support most input devices (e.g. scanners, cameras, word processors, paint/draw programs), supports a variety of data compression techniques, and has the flexibility to add new features easily in a controlled fashion [7,8,9]. TIFF files contain information on the height and width of the image, as well as resolution information, information about gray scale or color, and the type of compression scheme being used. TIFF is able to support resolutions of up to 300 dots per inch, and LaserROM took advantage of this by using TIFF to store high-resolution illustrations on the CD-ROM.

Selecting a Formatter

Once a document format had been decided upon, and the issue of illustrations in the manuals resolved, the next problem was determining how to take the documents in their SGML/HP Tag format and convert them to the format in which they would reside on the disc. This also meant defining the disc document format. This is discussed in the next section.

HP Tag is designed to allow the user the ability to specify the document structure using tags, and then executing a program which translates the tags into specific formatting instructions, producing a laser-printed manual with the appropriate design [5]. To perform this translation, Corporate Documentation Support provided a tool called MARKUP [4,6]. This is an SGML application generator which uses knowledge of the valid document structure to take a document in SGML format, and replace the tags with appropriate formatting commands, depending on what formatter is used. For instance, if a **<title>** tag were encountered, then MARKUP would know to replace this with formatting commands to the effect of centering the title on a new page, and putting it in boldface.

The LaserROM team, however, did not want to take the documents all the way to a laser printer, but to a compact disc. A customized version of MARKUP was

generated to replace the tags with TeX [10,11,12] formatting commands, so that the formatting power of TeX could be used to format the complex structures in the documentation.

TeX was developed by Donald Knuth at Stanford University. Because of the features it provides, its macro facility, and the fact that the HP Tag project at Corporate Documentation Support already was using TeX, this formatter was also chosen for LaserROM.

Design decisions (described below) dictated that the files on the compact disc would be in ASCII format. That meant formatting all the text in the tagged file and writing the result out to an ASCII flat file. Typeface changes would be controlled by control sequences, interpreted by the user interface at run time. Special information such as the date the manual was created, the title of the manual, and illustration information would also be described by control sequences. These control sequences would have to be transparent to the formatter. This was achieved using the **\special** function in TeX, which allows strings of characters to be passed through without formatting them. The TeX file which might be generated from Figure 2 is illustrated in Figure 3.

```
\special {←D}
\glossary
\title{Glossary of Animals}

\entry
\term{Aardvark}
\definition
The first animal listed in a dictionary.
\enddefinition
\endentry

\entry
\term{Dog}
\definition
A domesticated canine, probably descended
from the \xref{wolf}.
\enddefinition
\endentry
    .
    .
    .
\endglossary
```

Figure 3. The tags used to define the document structure in Figure 2 are replaced by corresponding TeX macro calls. TeX then interprets these command sequences as commands to perform formatting. Included in this formatting are the control sequences which will cause some text to be displayed in boldface, for instance.

Choosing TeX as the formatter meant that the output would be in device independent (DVI) format. Almost any typesetting device can be driven by a program that takes DVI files as input. This means the output of TeX is capable of being printed on many different kinds of equipment [11]. For LaserROM, the output device was actually an ASCII file. A DVI to ASCII driver was written to achieve this, taking into special consideration the control sequences as they were passed through.

Most importantly, a set of macros was written which would instruct TeX as to how to format the tagged input so that it would be output correctly in the ASCII file. These macros are responsible for taking into account the integer character widths and line spaces of an ASCII file, and insert the appropriate control sequences as needed to enhance the displayed text. They are very complex, as they must format tables, labeled lists, and syntax elements faithfully.

Data Representation

Defining the Physical Layout

Defining the disc format was really the resolution of two issues: the logical format of the documents that would reside on the disc so that they would integrate with the LaserROM interface; and defining the physical format of data. The physical layout had to conform to the "Yellow Book" CD-ROM specification, which was developed by Philips and Sony [13,14].

The High Sierra Group Proposal [1,13,14], which is the basis for International Standard 9660 (ISO 9660) [15], was adopted as the structure for the physical layout. This standard defines the volume and file structure for the data content on a CD-ROM. It is a read-only structure that ensures that all CD-ROMs may be read on virtually all combinations of CD-ROM playback units and computers. Together with the "Yellow Book", these two standards are generally considered the "CD-ROM Standard". This proposal was chosen because it was not proprietary, and was quickly becoming a standard within the CD-ROM community. The High Sierra format is currently used by firms such as Microsoft and Reference Technology, and the Microsoft CD-ROM extensions support this format under MS-DOS.

Defining the Logical Format

Defining the logical format was very much dependent upon the user interface, and its capabilities. The LaserROM interface required that the data be keyworded and indexed so that it could be retrieved quickly and effectively. The Ful/Text system, developed by Fulcrum Technologies, Inc. [16,17] was used as the text retrieval engine. Ful/Text is a full text retrieval system which includes tools for building a full custom interface. This retrieval system is designed to work with any collection of ASCII files on a variety of platforms. The LaserROM team simply extended this design by putting the ASCII files onto a CD-ROM.

The Ful/Text system provided tools to locate information on any topic in a collection of documents. The ASCII files generated by the DVI to ASCII conversion program were input to a series of programs which indexed them. On the retrieval end, the Ful/Text system searched its index files and retrieved documents relevant to any topic specified. The system also provided a specification of the document format [17], that is, how the ASCII file needed to be structured so that it could be indexed correctly by the software. This specification also included the definition of control sequences which would specify typeface changes. These were adapted for the LaserROM interface, and integrated into the TeX formatting step detailed in the previous section [18]. The TeX file in Figure 3 would appear much like Figure 4 when it had been processed through TeX and the DVI to ASCII driver.

The Ful/Text document format is based on the architecture defined in ISO's multilayered reference model of open systems interconnection. This document format is a specific data syntax for use by OSI presentation layer protocols (ISO 7498).

The Ful/Text process actually existed in two parts, analogous to the data preparation and interface portions of LaserROM.

←D¢32s
¢1mGlossary: Glossary of Animals¢21m

¢1mAardvark¢21m
The first animal in a dictionary.

¢1mDog¢21m
A domesticated canine, probably descended from the wolf.

.
.
.

Figure 4. This is an illustration of how the data is logically formatted for the CD-ROM. Embedded into the ASCII text are control sequences which allow the text to be indexed. Also shown are some control sequences which control font enhancements, such as boldface text. This file would be processed through the indexing programs before it is actually ready to go onto the compact disc.

The data preparation portion dealt with building the document collections, compressing and encrypting the data so that it was retrievable and readable only when using the proper interface, and indexing the documents. This was accomplished in a series of steps which processed the data from the point where it had been formatted by the TeX formatter and resided in an ASCII file (as in Figure 4), to the point where it was ready to be put into the High Sierra format. The interface portion is (discussed in detail in the next section), basically decompressed and decrypted the data on the CD-ROM, and provided the keyword retrieval engine for the LaserROM interface.

Final Disc Preparation

After the data had been indexed, compressed and built into the collections, and before the data was put into the High Sierra format, a final check was made of all the data as it would appear on the disc. Using the interface, the actual CD-ROM could even be simulated. This quality assurance step was used to proofread the documents for obvious errors, and to make sure that all the documents could be retrieved. This simulation was also important in estimating hardware dependencies such as seek times for different kinds of information. These seek times were directly affected by the order in which things were physically laid out on the compact disc, so, by trying different layouts of the data—for instance, putting all the illustrations at the end of the disc and all the manuals at the beginning of the disc—an optimum order for the physical layout could be determined.

To aid in the layout of the disc, and the simulation of the CD-ROM, a CD Publisher, built by Meridian Data, Inc., was purchased [19]. This machine allowed the development team to store the large amounts of information being processed on its internal disks, as well as providing the utilities to do actual emulations of different manufacturer's CD-ROM drives. The CD-Publisher came with software which would convert the MS-DOS files generated by the indexing engine, and convert them to High Sierra format. Finally, it provided the capability to create the master tapes required by the compact disc mastering facilities, in the variety of formats required by the different compact disc manufacturers. The advantage of processing the data to this point was that the LaserROM project was able to control the CD-ROM development process up to the point of actually producing the CD-ROM. The only point where any other vendors were directly involved was in the actual compact disc production process.

The LaserROM Interface

Choosing a Retrieval System

The LaserROM project was designed to run on IBM-PC compatible machines. This meant providing a user interface which would integrate nicely into the PC office workstation environment. This also meant developing an interface which could be distributed easily and widely to the PC users. The keystone of the LaserROM interface was the document retrieval system.

The Ful/Text full text retrieval system discussed in the previous section was also chosen because it provided a toolkit with which to customize a user interface. It was on this platform that the LaserROM interface was developed. The Ful/Text retrieval system fully implemented many of the classic text retrieval features [20], including: phrase and proximity searching; thesaurus capabilities; expansion or "wildcard" searching; boolean searching; and date range and field searches [2,21]. These features provided a rich set of search capabilities which LaserROM was able to exploit. Using the search primitives provided, the next step was to determine how the actual interface would appear to the user, that is, the actual "look and feel" of the LaserROM interface.

Choosing an Interface

In keeping with the approach taken in choosing the Ful/Text retrieval system, the development team explored what other vendors provided in the way of a development environment for user interfaces. This concept is often referred to as the "toolkit" approach. The underlying search and retrieval software was provided, and the LaserROM developers looked for a toolkit that would provide the ability to construct an interface on top of this basic "search engine." This would provide the developers with full customization of the portion that the user accesses, without the total development time that would be required to build an entire system [2].

Microsoft Windows was chosen as the basis for the interface. There were a few factors affecting this decision. First, Windows provided the user with an intuitive, easy-to-use interface. Windows was a proven technology, and was fast becoming a standard amongst the users of personal computers. Utilizing the toolkit approach, the LaserROM developers could decrease their time-to-market. Finally, MS-Windows was already being used in several applications throughout the corporation, so that LaserROM would have the "look and feel" of other HP PC applications.

After the user interface had been designed, usability tests were conducted to make sure that people could use the product, and this feedback was used to improve the graphical interface.

LaserROM was introduced in September, 1987. The first subscription disc was distributed in April, 1988. By taking advantage of existing tools and standards, the time-to-market for the project was relatively short. The LaserROM service is a subscription service which provides monthly updates of user manuals and support information. By integrating the tools and standards into the data preparation process, the CD-ROM production process has become virtually automated, and a number of subscription discs can be produced on a monthly basis. This provides the customers with current information that is immediately at their disposal within the framework of the PC workstation environment.

CD-ROM Issues and the Future of LaserROM

CD-ROM Issues

Choosing the CD-ROM technology as a vehicle for distributing Hewlett-Packard's user manuals and support information presented some important hardware and software issues that needed to be dealt with. Since LaserROM was to run on an IBM-PC compatible environment, this meant making use of Microsoft's MS-DOS CD-ROM extensions to be able to recognize and access a CD-ROM drive [22]. This software fit nicely with the use of MS-Windows as an interface.

Next, a CD-ROM drive needed to be chosen. At the time, there were a number of drives on the market which were external to the PC, and connected by an interface cable and a controller card. However, a number of vendors were already developing half-height CD-ROM drives which could be installed directly into one of the floppy-disc slots on a PC. A number of both the external and internal drives were tested, and it was decided that the half-height drives would be used. HP became the first PC manufacturer to offer the internal half-height CD-ROM drives. These were manufactured by Philips, Inc., and provided a fully integrated solution for LaserROM.

The Future of LaserROM

With the use of technologies such as CD-I or DV-I (Compact Disc- or Digital Video-Interactive), full-motion video on a CD-ROM disc is now possible. LaserROM could theoretically contain interactive manuals that would guide the user with both text and video as to how to fix a disk drive head crash, complete with an audio narration to take the user through the sequence step-by-step. The LaserROM development

team is actively pursuing avenues of growth in the area of mixed-media.

Another area of opportunity for LaserROM is in the field of hypermedia. Tags already exist within HP Tag which could provide LaserROM with a limited hypertext capability. These cross-reference tags could be the basis for the linkages that would form a natural "knowledge path" that might be followed when searching or browsing the data [1,2].

Conclusion

The LaserROM project encompasses many of the issues that are being addressed by document processing systems today. The entire product was conceived as a vehicle for disseminating large amounts of information to Hewlett-Packard's customers and make that information more useful. Out of this grew the easy-to-use user interface, with its keyword retrieval system to aid in the search for relevant information. To get to this point, however, an entire data preparation process had to be designed, and this was accomplished by integrating existing tools and standards into the process. This process, along with the user interface and the actual CD-ROM, all work together to provide a working example of a successful document processing system.

References

[1] Ropiequet, Suzanne (ed.), Einberger, John, and Zoellieck, Bill. *CD-ROM: Optical Publishing.* Microsoft Press, Redmond, Wash., 1987.

[2] Ferguson, Greg. "User Interface Design Methodologies for CD-ROM Information Retrieval." 1988 INTEREX Proceedings, Orlando, Fla., Aug. 1988.

[3] International Organization for Standardization (ISO). *ISO 8879-1986(E). Information Processing - Text and office systems - Standardized Generalized Markup Language (SGML).* International Organization for Standardization, Switzerland, 1986.

[4] Hewlett-Packard Corporate Documentation Support. *MARKUP Reference Manual.* Hewlett-Packard, Palo Alto, Calif., Mar. 1988. Internal documentation.

[5] Hewlett-Packard Corporate Documentation Support. *HP Tag/Vectra User's Guide.* Hewlett-Packard, Palo Alto, Calif., Nov. 1987. Internal documentation.

[6] Price, Lynne and Schneider, Joe. "Evolution of an SGML Application Generator." *Proceedings, ACM Conference on Document Processing Systems,* Santa Fe, N.M., Dec., 1988.

[7] Andrews, Nancy and Fry, Stan. "TIFF: An Emerging Standard for Exchanging Digitized Graphics Images." *Microsoft Systems Journal.* July, 1987.

[8] Aldus Corporation and Microsoft Corporation. *Tag Image File Format—Rev 4.0.* An Aldus/Microsoft technical memorandum, April, 1987.

[9] Meadow, Anthony and Offner, Rocky and Budiansky, Michael. "Handling Image Files with TIFF." *Dr. Dobb's Journal,* May, 1988.

[10] Knuth, Donald E. *The TeXbook.* Addison Wesley, Reading, Mass., 1986.

[11] Knuth, Donald E. *TeX: The Program.* Addison Wesley, Reading, Mass., 1986.

[12] Vose, G. Micheal and Williams, Gregg. "Computer Science Considerations." *Byte.* February, 1986.

[13] Buddine, Laura and Young, Elizabeth. *The Brady Guide to CD-ROM.* Prentice Hall, New York, 1987.

[14] Schwerin, Julie B. *CD-ROM Standards: The Book.* Learned Information, Ltd. and InfoTech, Oxford, England, 1986.

[15] National Information Standards Organization. *ANS Z39.60-198X. Proposed American National Standard for Information Sciences Volume and File Structure of CD ROM for Information Interchange.* National Bureau of Standards, Gaithersburg, Maryland.

[16] Fulcrum Technologies, Inc. *Ful/Text Reference Manual, Version 4.2.* Fulcrum Technologies, Inc., Ottawa, Canada, 1987.

[17] Fulcrum Technologies, Inc. *Ful/Text Programmer's Guide, Version 4.2.* Fulcrum Technologies, Inc., Ottawa, Canada, 1987.

[18] Miller, James R. "LaserROM Raw Database Specification." Hewlett-Packard internal documentation, 1987.

[19] Meridian Data, Inc. *CD Publisher User Manual.* Meridian Data, Inc. Capitola, Calif., 1988.

[20] Salton, Gerard, and McGill, M. J. *Introduction to Modern Information Retrieval.* McGraw-Hill, New York, 1983.

[21] Hewlett-Packard. *LaserROM Quick Reference Guide.* Part No. 5959-5411. Hewlett-Packard, Mountain View, Calif., 1988.

[22] Microsoft, Corp. *Microsoft MS-DOS CD ROM Extensions.* Microsoft, Corp. Redmond,Wash., 1988.

Other useful references.

Lambert, Steve and Ropiequet, Suzanne (ed.). CD-ROM: *The New Papyrus*. Microsoft Press, Redmond, Wash., 1987.

Roth, Judith Paris (ed.) *Essential Guide to CD-ROM.* Meckler Publishing, Westport, Conn., 1986.